Book Review

Mining the Social Web: Analyzing Data from Facebook, Twitter, LinkedIn, and Other Social Media Sites

Social Media Concepts in Action for Practitioners and Hobbyists Alike,

Mining the Social Web by Matthew Russell is a delightful reading which covers a wide array of very pertinent social media technologies. When working in the field of social media and web mining, one comes across a plethora of buzz words, tools and techniques to do various web data analysis tasks. This book does a great job at demystifying these topics. What I liked most about Mining the Social Web is that it takes a specific area such as TF-IDF and then connects the dots with various intertwined technologies (bi-grams, graph visualization, cosine similarity) to paint the big picture in terms of practical usage. Even though there are several prerequisites (python is a big one) and a small learning curve for the tools introduced in the book, it's well written and easy to follow.

Web data analysis is quite a huge field of research and development; given it's hard to cover a majority of topics in one book with practical examples, author has done a tremendous job by addressing the practical and relevant concerns such as accessing the twitter field with oAuth, mapReduce, NLP, couchDB, processing mailboxes, accessing linkedin and last but not least the virtues of semantic web. Nowadays when most technical books usually tend to become occasional reference books where you refer to specific topic and skim through the rest, mining the social web is a big exception. Either you are trying to familiarize yourself with the current technologies of social media mining, or are a experienced developer looking for a different point of view for implementation, this book will definitely widen your horizons.

For book reviews, please feel free to contact Adnan Masood at adnan@nova.edu
Adnan Masood works as a web architect and technical lead for Green Dot Corporation where he develops SOA based middle-tier architectures, distributed systems, and web-applications using Microsoft technologies. He is a Microsoft Certified Trainer holding several technical certifications, including MCPD (Enterprise Developer), MCSD .NET, and SCJP-II. Adnan is attributed and published in print media and on the Web; he is technical editor for "Microsoft Windows Server AppFabric Cookbook" and also taught Windows Communication Foundation (WCF) courses at the University of California at San Diego.

Adnan regularly presents at local code camps and user groups. He is actively involved in the .NET community as cofounder and president of the of San Gabriel Valley .NET Developers group. Adnan holds a Master’s degree in Computer Science; he is currently a doctoral student working towards PhD in Machine Learning; specifically interestingness measures in outliers using Bayesian Belief Networks. He also holds systems architecture certification from MIT and SOA Smarts certification from Carnegie Mellon University.